Homework

MATH285K - Spring 2010 *Instructor: Sebastien Roch*

Instructions

This optional homework is due on **Friday June 4 in class**. Late homeworks will not be accepted. The assignment should be **typed** and is to be done **without consultation**.

Questions

1. Let $p(n, f)$ denote the number of (not necessarily binary) phylogenetic trees with label set $X = \{1, \ldots, n\}$ and f interior edges.

a) For all $n \geq 3$, establish the recursion

$$p(n, f) = (f + 1)p(n - 1, f) + (n + f - 2)p(n - 1, f - 1).$$

b) Using this recursion, deduce that the number of binary phylogenetic trees on n leaves is

$$b(n) = (2n - 5)!! \equiv (2n - 5) \times (2n - 7) \times \cdots \times 3 \times 1.$$

2. Let $\mathcal{T}$ be a binary phylogenetic X-tree with $|X| = 2m$.

a) Show that in $\mathcal{T}$ there is a unique set of m edge-disjoint paths, each of which joins two leaves of $\mathcal{T}$.

b) Deduce that the number of binary characters $\chi : X \to \{0, 1\}$ with parsimony score m on $\mathcal{T}$ is exactly 2^m.

3. An X-tree $\mathcal{T}$ is said to *display a partial split* $A|B$ *of* X if $A|B$ is a split of $\mathcal{T}|(A \cup B)$. Suppose that an X-tree displays the partial splits $A_1|B_1$ and $A_2|B_2$ of X, where $A_1 \cap A_2$, $A_1 \cap B_2$, and $B_1 \cap B_2$ are all non-empty. Show that $\mathcal{T}$ also displays the partial splits $(A_1 \cup A_2)|B_1$ and $A_2|(B_1 \cup B_2)$ of X.

4. Let (X, d) be a metric space with $|X| = 3$. Show that d is a tree metric on X.

5. Let $\mathcal{T} = (T, \phi)$ be a phylogenetic X-tree, $w : E \to \mathbb{R}_+$ a real-valued weighting of the edges of T, and d the corresponding tree metric on X. Let $W = \sum_e w_e$. For every pair of distinct $x, y \in X$, let $I(x, y)$ denote the set of interior vertices of T in the path connecting $\phi(x)$ and $\phi(y)$. Let $\lambda : X \times X \to \mathbb{R}_+$ be a dissimilarity map on X defined, for all $x, y \in X$, in terms of degrees of the vertices of T (i.e., $D(v)$ is the degree of v) as follows:

$$\lambda(x, y) = \begin{cases} \prod_{v \in I(x,y)} (D(v) - 1)^{-1}, & \text{if } x \neq y, \\ 0, & \text{if } x = y. \end{cases}$$

Show that

$$W = \sum_{\{x,y\} \subseteq X} \lambda(x, y) d(x, y).$$

6. Consider the CFN model with state space $C = \{0, 1\}$ on a phylogenetic X-tree and a character $\chi : X \to C$. Show that

$$\mathbb{P}[\chi] \leq \frac{1}{2} p_{\max}^{\ell(\chi, \mathcal{T})},$$

where $\ell(\chi, \mathcal{T})$ is the parsimony score of χ on $\mathcal{T}$ and

$$p_{\max} = \max\{p(x, y) : x, y \in X,\ \chi(x) \neq \chi(y)\},$$

with $p(x, y)$, the probability that the states at $\phi(x)$ and $\phi(y)$ disagree.

7. Consider a GTR model on a phylogenetic X-tree with rate matrix Q, stationary distribution π, and tree metric $\kappa(x, y)$. Show that

$$p(x, y) = 1 - \operatorname{tr}\left(\Pi \exp\left(Q \frac{-\kappa(x, y)}{\operatorname{tr}(\Pi Q)}\right)\right),$$

where Π is the diagonal matrix with diagonal π.